Daily Report

Yuxin Ma

09.09.2013

Research

Tours Method Applied to SVM Visualization

Today there is a brand new idea for visualizing SVM decision boundary and data instances based on Tours method.

Tours method[1, 2] is a linear combinations (projections) of original real-valued d dimensional space. Columns in the transformation matrix $\boldsymbol{A} \in \mathbb{R}^{d \times p}$ is orthonormal while p is usually between 1 and 3 to transform the original data into a visual space. $\boldsymbol{A}$ can be set freely to display projection from any direction. Besides linear combinations, the tours method introduces "path" of projections. Take $p = 2$ (i.e., projection onto a 2-D plane) into consideration, the path is a trajectory of a continuous series of projection planes, which enables users to explore projection patterns as if they are "traveling" in the data.

[3, 4, 5] applied tours method to SVM visualization. The author constructed a projection around the separating plane to show the distribution of data and decision boundaries in binary classification configuration.

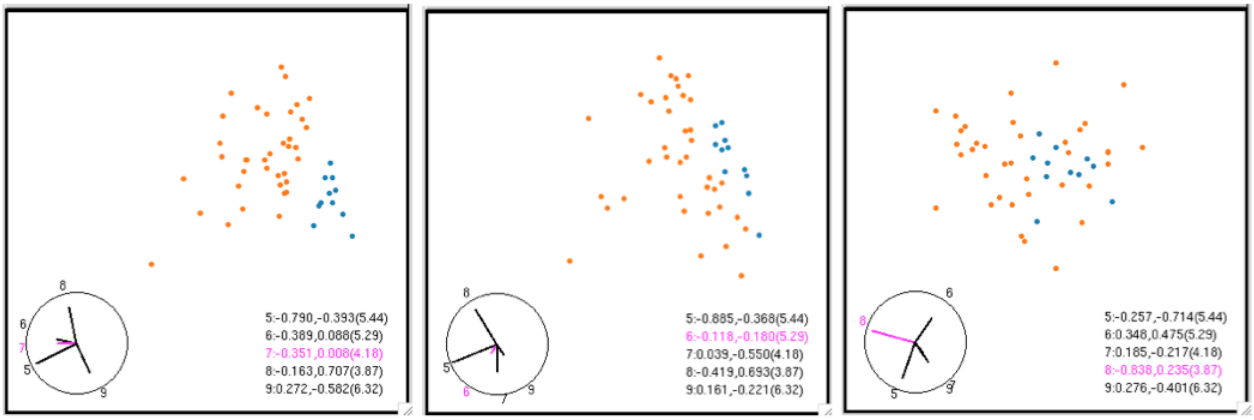

Figure 1: Example of tours method applied on a data with 5 dimensions from [4]. Under different $\boldsymbol{A}$, the three figures shows the linear decision boundary from different aspect, which seems like a slight continuous "clockwise vertical rotation" from right to left.

What we can do further In these papers, the tours method is just a displaying method with static figures. Based on the tours method, we can perform more interactions for users to view the distribution, the classification result and guide them to optimize parameters. Extraction of classification rules can be more convenient in this view, for it is a linear transformation

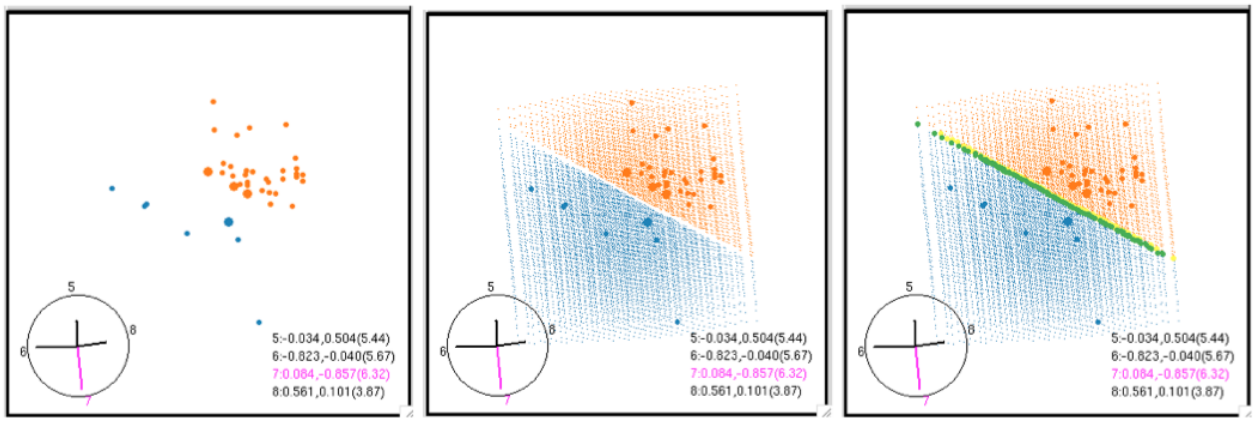

Figure 2: Classification result of a 4-dimension dataset involving decision boundary from [4]. Left: 2 classes of data marked by two different color. Larger dots indicate support vectors. Middle: a grid of sample points colored by their predicted label. The gap indicates position that has predicted value close to 0, which is shown as green and yellow dots in the right figure.

and easier to find the mapping of selected data instances in the view from the projection view to their original dimensions.

Plans

Today I will do more research on the tours method and try to figure out a way to show the details around decision boundaries in a more interactive manner beyond what the existed papers have already done.

References

[1] C.-h. Chen, W. Härdle, and A. Unwin, *Handbook of Data Visualization*. Springer, Dec. 2007.

[2] A. Buja, D. Cook, D. Asimov, and C. Hurley, "Computational methods for high-dimensional rotations in data visualization," *Handbook of statistics: Data mining and data visualization*, vol. 24, pp. 391–413, 2005.

[3] D. Caragea, D. Cook, and V. G. Honavar, "Gaining insights into support vector machine pattern classifiers using projection-based tour methods," in *KDD '01: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM Request Permissions, Aug. 2001.

[4] D. Caragea, D. Cook, H. Wickham, and V. Honavar, "Visual Methods for Examining SVM Classifiers," *Visual Data Mining*, Apr. 2008.

[5] D. Caragea, D. Cook, and V. Honavar, "Visual methods for examining support vector machines results, with applications to gene expression data analysis," tech. rep., 2005.